8 Lecture 8

8.1 Existence, uniqueness and basic properties of the determinant

Theorem 8.1. *There is a unique function* $\det : M_{n\times n}(F) \to F$ *satisfying the properties in Definition 7.12. In fact, the determinant is given by*

$$\det(A) = \sum_{\sigma\in S_n} \operatorname{sign}(\sigma) a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n}.$$

This is called the ***Leibniz formula for determinant****.*

Proof. Let $e_1, \dots, e_n$ be the standard basis of F^n. By multilinearity,

$$\begin{aligned} \det(A) &= \det(\sum_{i_1=1}^{n} a_{i_1 1} e_i \ \dots \ \sum_{i_n=1}^{n} a_{i_n n} e_i) \\ &= \sum_{i_1=1}^{n} \cdots \sum_{i_n=1}^{n} a_{i_1 1} a_{i_2 2} \dots a_{i_n n} \det(e_{i_1} \ \dots \ e_{i_n}) \end{aligned}$$

By the alternating property, $\det(e_{i_1} \ \dots \ e_{i_n}) = 0$ if $i_k = i_l$ for $k \neq l$. Thus the summation is over all $i_1, \dots, i_n$ which are distinct. We would like to permute $e_{i_1} \ \dots \ e_{i_n}$ to $e_1, \dots, e_n$ and use the fact that $\det I = 1$. By Lemma 7.13, each swap will alter the sign of det. The question is how many swaps do we need to permute $e_{i_1} \ \dots \ e_{i_n}$ to $e_1, \dots, e_n$?

Definition 8.2. A **permutation of** n **elements** is a bijective map $\sigma : \{1, \dots, n\} \to \{1, \dots, n\}$. To write a permutation, we just list all its values $\sigma(1)$ to $\sigma(n)$. We use the one-line notation to write $\sigma = \sigma(1)\cdots\sigma(n)$. The set of all permutation is denoted by S_n which is called the **symmetric group**. There are $n! = n(n-1)\cdots 1$ elements in S_n.

An **inversion** in σ is a pair of indices (i, j) such that $i < j$ and $\sigma(i) > \sigma(j)$.

The **inversion number** of σ, denoted $\operatorname{inv}(\sigma)$, is the total number of inversions in σ.

The **sign** of a permutation σ is $\operatorname{sign}(\sigma) = (-1)^{\operatorname{inv}(\sigma)}$.

Example 8.3. $\operatorname{inv}(123) = 0$, $\operatorname{inv}(231) = 2$, $\operatorname{inv}(312) = 2$, $\operatorname{inv}(213) = 1$, $\operatorname{inv}(132) = 1$, $\operatorname{inv}(321) = 3$.

Lemma 8.4. *The number of adjacent swaps needed to sort* σ *into the identity permutation in the bubble sort algorithm is* $\operatorname{inv}(\sigma)$*. In particular,* $\det(e_{\sigma(1)} \ \dots \ e_{\sigma(n)}) = \operatorname{sign}(\sigma)$.

Proof. Given a list $a_1, a_2, \dots, a_n$, the bubble sort proceeds as follows:

(i) For each $1 \le i \le n-1$, compare a_i and a_{i+1}.

(ii) If $a_i > a_{i+1}$, swap them. If $a_i < a_{i+1}$, keep them unchanged.

(iii) After one full pass through the list, the largest element "bubbles up" to the end of the list.

(iv) Repeat the above passes on the first $n-1$, then $n-2$, …, elements, until no swaps occur during a pass.

Consider bubble sort applied to the sequence $\sigma(1), \dots, \sigma(n)$. We claim that each swap removes exactly one inversion. Bubble sort only swaps adjacent elements $\sigma(k)$ and

$\sigma(k+1)$ when $\sigma(k) > \sigma(k+1)$. This removes the inversion $(k, k+1)$. We now verify that the number of other inversions remains the same. Let $(i, j) \neq (k, k+1)$ be an inversion before the swap. Consider the following cases:

If $i < k$ and $j > k+1$, then $\sigma(i)$ and $\sigma(j)$ are unchanged — the inversion remains.
If $i < k$ and $j = k$, the inversion becomes $(i, k+1)$ after the swap.
If $i < k$ and $j = k+1$, it becomes (i, k).
If $i = k$ and $j > k+1$, it becomes $(k+1, j)$.
If $i = k+1$ and $j > k+1$, it becomes (k, j).

Therefore, each swap in bubble sort decreases the total number of inversions by exactly one. Since bubble sort terminates when the sequence is sorted (i.e., has zero inversions), the total number of swaps performed is equal to the number of inversions in the original sequence. $\square$

Coming back to the proof since $i_1, \ldots, i_n$ are distinct, there is $\sigma \in S_n$ such that $i_k = \sigma(k)$ for all k. Therefore

$$\begin{aligned}\det(A) &= \sum_{\sigma \in S_n} a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n} \det(e_{\sigma(1)} \ \ldots \ e_{\sigma(n)}) \\ &= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n}.\end{aligned}$$

The above argument show that any function satisfying Definition 7.12 has to have the above expression. We now check that the function given by the expression above does satisfies Definition 7.12.

(i) Fix all columns except column j. Suppose $v_j = ax + by$. Then each entry in column j satisfies $a_{ij} = ax_i + by_i$ for all $i = 1, \ldots, n$.

$$\begin{aligned}&\sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(j)j} \cdots a_{\sigma(n)n} \\ =&\sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{\sigma(1)1} \cdots (ax_{\sigma(j)} + by_{\sigma(j)}) \cdots a_{\sigma(n)n} \\ =&a \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{\sigma(1)1} \cdots x_{\sigma(j)} \cdots a_{\sigma(n)n} + b \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{\sigma(1)1} \cdots y_{\sigma(j)} \cdots a_{\sigma(n)n}.\end{aligned}$$

(ii) Suppose two columns are equal: $v_i = v_j$ for $i \neq j$. For any $\sigma \in S_n$ let $\sigma' \in S_n$ such that $\sigma'(i) = \sigma(j)$ and $\sigma'(j) = \sigma(i)$ and $\sigma'(k) = \sigma(k)$ for $k \neq i, j$. We have seen in the proof of Lemma 8.4 that swapping two adjacent elements in a permutation change the inversion by ± 1. Since we need $2|j-i|+1$ adjacent swaps to permute σ' to σ we have $\operatorname{sign}(\sigma') = (-1)^{2|j-i|+1} \operatorname{sign}(\sigma) = -\operatorname{sign}(\sigma)$. Since columns i and j are equal, $a_{\sigma(1)1} \cdots a_{\sigma(n)n} = a_{\sigma'(1)1} \cdots a_{\sigma'(n)n}$. Thus we have

$$\begin{aligned}&\operatorname{sign}(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n)n} + \operatorname{sign}(\sigma') a_{\sigma'(1)1} \cdots a_{\sigma'(n)n} \\ =&(\operatorname{sign}(\sigma) + \operatorname{sign}(\sigma')) a_{\sigma(1)1} \cdots a_{\sigma(n)n} = 0.\end{aligned}$$

Thus the terms corresponding to σ and σ' cancel in the sum. As every permutation σ corresponds to exactly one such pair, we have $\sum_{\sigma\in S_n} \operatorname{sign}(\sigma)a_{\sigma(1)1}a_{\sigma(2)2}\cdots a_{\sigma(n)n} = 0$.

(iii) Recall the identity matrix $I = (\delta_{ij})$, so we have

$$\delta_{\sigma(1)1}\delta_{\sigma(2)2}\cdots\delta_{\sigma(n)n} = \begin{cases} 1 & \text{if } \sigma = \text{id} \\ 0 & \text{otherwise} \end{cases}$$

Only the identity permutation contributes, and $\operatorname{sign}(\text{id}) = 1$, so:

$$\sum_{\sigma\in S_n} \operatorname{sign}(\sigma)\delta_{\sigma(1)1}\delta_{\sigma(2)2}\cdots\delta_{\sigma(n)n} = 1.$$ □

Example 8.5. $\det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22}\end{pmatrix} = \operatorname{sign}(12)a_{11}a_{22} + \operatorname{sign}(21)a_{21}a_{12} = a_{11}a_{22} - a_{21}a_{12}$.

$$\begin{aligned}\det\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33}\end{pmatrix} =& \operatorname{sign}(123)a_{11}a_{22}a_{33} + \operatorname{sign}(231)a_{21}a_{31}a_{13} + \operatorname{sign}(312)a_{31}a_{12}a_{23} \\ &+ \operatorname{sign}(213)a_{21}a_{12}a_{33} + \operatorname{sign}(132)a_{11}a_{32}a_{23} + \operatorname{sign}(321)a_{31}a_{22}a_{13} \\ =& a_{11}a_{22}a_{33} + a_{21}a_{31}a_{13} + a_{31}a_{12}a_{23} - a_{21}a_{12}a_{33} - a_{11}a_{32}a_{23} - a_{31}a_{22}a_{13}\end{aligned}$$

The Leibniz formula is of theoretical importance but it's not a good idea to compute the determinant using this formula for large matrices.

Corollary 8.6.

$$\det(P_{ij}) = -1$$
$$\det(D_i(\lambda)) = \lambda$$
$$\det(E_{ij}(\lambda)) = 1$$

For any $A \in M_{n\times n}(F)$ and elementary matrix E

$$\det(AE) = \det(A)\det(E)$$

Proof. Since P_{ij} is obtained by swaping column i and j of I, we have $\det(P_{ij}) = -1$.

Since $D_i(\lambda)$ is obtained by multiplying column j of I by λ, we have $\det(D_i(\lambda)) = \lambda$.

Since $E_{ij}(\lambda)$ is obtained by adding $\lambda\cdot$ column i to column j of I, we have $\det(E_{ij}(\lambda)) = 1$.

By Remark 6.23, multiplying elementary matrices to the right of a matrix corresponds to performing elementary column operations. Since each elementary column operation corresponds to multiplying the determinant of A by $\det(E)$, we have $\det(AE) = \det(A)\det(E)$. □

Corollary 8.7. *Let $A \in M_{n\times n}(F)$ be singular. Then* $\det(A) = 0$.

Proof. Since $\operatorname{rank} A < n$, the columns of A are linearly dependent. We write $A = (v_1 \dots v_n)$. By Lemma 2.8, there is a column v_k which is a linear combination of $v_1, \dots, v_{k-1}$ i.e. $v_k = a_1 v_1 + \cdots + a_{k-1} v_{k-1}$. Thus by Lemma 7.13, adding $-a_1 v_1 - \cdots - a_{k-1} v_{k-1}$ to v_k we get $\det(A) = \det(v_1 \dots v_{k-1}, 0, v_{k+1} \dots v_n) = 0$. □

Proposition 8.8. *Let* $A, B \in M_{n\times n}(F)$. *Then* $\det(AB) = \det(A)\det(B)$.

Proof. In case B is not invertible, we have first that $\det(B) = 0$. Next, we know that AB is not invertible by Homework 4 (ST invertible $\Longrightarrow S, T$ are both invertible). Hence $\det(AB) = 0 = \det(A)\det(B)$.

We have seen in Corollary 8.6 that $\det(AE) = \det(A)\det(E)$ for any $A \in M_{n\times n}(F)$ and elementary matrix E. By Corollary 7.9 for any invertible matrix B, we have $B = E_1 \cdots E_m$ for some elementary matrices $E_1, \dots, E_m$. Hence

$$\begin{aligned}\det(AB) &= \det(AE_1 \cdots E_{m-1}E_m) = \det(AE_1 \cdots E_{m-1})\det(E_m)\\ &= \cdots\\ &= \det(A)\det(E_1)\cdots\det(E_m)\\ &= \det(A)\det(E_1 \cdots E_m) = \det(A)\det(B).\end{aligned}$$

□

Corollary 8.9. $A \in M_{n\times n}(F)$ *is invertible if and only if* $\det(A) \neq 0$. *If A is invertible, then* $\det(A)^{-1} = \frac{1}{\det(A)}$.

Proof. If A is not invertible then by Corollary 8.7, $\det(A) = 0$.

If A is invertible Then $A^{-1}A = I$. By Proposition 8.8, $\det(A)\det(A^{-1}) = \det(AA^{-1}) = \det I = 1$. Thus $\det(A) \neq 0$ and $\det(A^{-1}) = \frac{1}{\det(A)}$. □

Remark 8.10. This Corollary shows that computing the determinant is a good way to tell whether a matrix is invertible or not.

Corollary 8.11. *If* $B = P^{-1}AP$, *then* $\det(B) = \det(A)$. *The determinant is independent of the choice of a basis. That is if* $T : V \to V$ *is linear then for any bases* β *and* γ *of* V, *we have* $\det([T]_\beta^\beta) = \det([T]_\gamma^\gamma)$. *Thus we define* $\det(T) = \det([T]_\beta^\beta)$.

Proposition 8.12. *Let* $A \in M_{n\times n}(F)$. *Then* $\det(A^t) = \det(A)$.

Proof. If A is not invertible, then by Corollary 6.17 A^t is not invertible. Thus $\det(A^t) = 0 = \det(A)$.

If A is invertible then by Corollary 7.9, $A = E_1 \cdots E_m$ for some elementary matrices $E_1, \dots, E_m$. Note that for any elementary matrix E, E^t is also an elementary matrix and $\det(E) = \det(E^t)$. Then

$$\begin{aligned}\det(A) = \det(E_1)\cdots\det(E_m) = \det(E_m^t)\cdots\det(E_1^t)\\ = \det(E_m^t \cdots E_1^t) = \det((E_1 \cdots E_m)^t) = \det(A^t).\end{aligned}$$

□

Remark 8.13. Thus the properties in Definition 7.12 and Lemma 7.13 are also true for rows of A.

8.2 Laplace expansion

Proposition 8.14. *Let $B \in M_{(n-1)\times(n-1)}(F)$ and $u \in F^{n-1}$. Then*

$$\det \begin{pmatrix} 1 & u^t \\ 0 & B \end{pmatrix} = \det B.$$

In particular

$$\det \begin{pmatrix} \lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} = \lambda_1 \dots \lambda_n.$$

Proof. By Lemma 7.13, we can perform elementary column operations to eliminate the u^t using the first column. For instance one first add $-u_1$ multiple of column 1 to column 2 to eliminate u_1 and then add $-u_2$ multiple of column 1 to column 3 to eliminate u_2 ... so on and so forth and finally add $-u_{n-1}$ multiple of column 1 to column n to eliminate u_{n-1}. We can express this process compactly as elementary column operation by block: one add $-u^t$ multiple of the first column $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ to the column $\begin{pmatrix} u^t \\ B \end{pmatrix}$ to get $\begin{pmatrix} 0 \\ B \end{pmatrix}$ which is equivalent to the matrix multiplication $\begin{pmatrix} 1 & u^t \\ 0 & B \end{pmatrix}\begin{pmatrix} 1 & -u^t \\ 0 & I \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & B \end{pmatrix}$.

Since the determinant is invariant under add multiples of other columns to one column, we have $\det \begin{pmatrix} 1 & u^t \\ 0 & B \end{pmatrix} = \det \begin{pmatrix} 1 & 0 \\ 0 & B \end{pmatrix}$. We define $D(B) = \det \begin{pmatrix} 1 & 0 \\ 0 & B \end{pmatrix}$. Clearly, this function satisfies the properties in Definition 7.12 since det does. Therefore, by the uniqueness of the determinant function, it must be $D(B) = \det(B)$.

We apply the previous result on inductively and get

$$\begin{aligned} \det \begin{pmatrix} \lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} &= \lambda_1 \det \begin{pmatrix} 1 & * & \cdots & * \\ 0 & \lambda_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \\ =& \lambda_1 \det \begin{pmatrix} \lambda_2 & \cdots & * \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix} = \cdots = \lambda_1 \dots \lambda_n. \end{aligned}$$

$\square$

Remark 8.15. A good way of computing the determinant is to use elementary row and column operations to reduce the matrix to the upper triangular form and apply the previous proposition.

Remark 8.16. In the proof above, we eliminate the block u^t as if it is a number that can be eliminated by the 1 in the same row. This is an efficient way to perform elementary row/column operations called elementary row/column operations by blocks. A general form of it is as follows.

Swap two block rows

$$\begin{pmatrix} A_{21} & A_{22} \\ A_{11} & A_{11} \end{pmatrix} = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}.$$

Multiply a block row by a square matrix Λ from the left:

$$\begin{pmatrix} \Lambda A_{11} & \Lambda A_{12} \\ A_{21} & A_{22} \end{pmatrix} = \begin{pmatrix} \Lambda & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}.$$

Add a block multiple of one block row to another

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} + MA_{11} & A_{22} + MA_{12} \end{pmatrix} = \begin{pmatrix} I & 0 \\ M & I \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

where M is a matrix of appropriate size (not necessarily invertible). This generalizes the usual row addition $R_i \leftarrow R_i + cR_j$ to the block setting.

The third one is most useful in computing determinants since the determinant is unchanged under such operations. The first two will differ by the determinant of the corresponding block elementary matrix.

Similarly, one can also perform elementary column operations by block. In particular adding a block multiple of one block column to another column is as follows

$$\begin{pmatrix} A_{11} & A_{12} + A_{11}M \\ A_{21} & A_{22} + A_{21}M \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} I & M \\ 0 & I \end{pmatrix}$$

where M is a matrix of appropriate size (not necessarily invertible).

We note that when doing the third elementary column operations by blocks, one can only multiply M to the right of a block column while when doing elementary row operations by blocks, one can only multiply M to the left of a block row.

Definition 8.17. The (i, j)**-minor** of A, denoted by A_{ij} is the determinant of the $(n-1) \times (n-1)$ matrix obtained by deleting the i-th row and the j-th column of A. The **cofactor matrix** $\mathrm{Cof}(A)$ is the matrix with entries $\mathrm{Cof}(A)_{ij} = (-1)^{i+j} A_{ij}$.

Example 8.18. If $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ then $\mathrm{Cof}(A) = \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}$.

Theorem 8.19 (Laplace's expansion formula)**.** *Let $A \in M_{n\times n}(F)$. Then for any $j = 1, \ldots, n$*

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} A_{ij} = \sum_{i=1}^{n} (-1)^{i+j} a_{ji} A_{ji}.$$

Proof. Let $e_1, \dots, e_n$ be the standard basis of F^n. By multilinearity,

$$\begin{aligned}\det(A) &= \det(v_1 \, \dots \sum_{i=1}^{n} a_{ij} e_i \, \dots \, v_n) \\ &= \sum_{i=1}^{n} a_{ij} \det(v_1 \, \dots \, e_i \, \dots \, v_n)\end{aligned}$$

where e_i is in the j-th column. We first do adjacent column swaps to make e_i appear in the first column and then do adjacent row swaps to make the entry 1 in e_i appear in the first row. Then we made $i-1+j-1 = i+j-2$ adjacent swaps and hence the determinant change by a factor $(-1)^{i+j-2} = (-1)^{i+j}$. We apply Proposition 8.14 to see that $\det(v_1 \, \dots \, e_i \, \dots \, v_n) = (-1)^{i+j} A_{ij}$

The second equality follows from applying the formula to A^t and use Proposition 8.12. $\square$

Definition 8.20. For $A = (v_1 \, \dots \, v_n) \in M_{n\times n}(F)$ define $A_i(b) = (v_1 \, \dots \, v_{i-1} \, b \, v_{i+1} \, \dots \, v_n)$ as the matrix obtained by replacing the i-th column of A with the vector $b \in F^n$.

Lemma 8.21.

$$\det(A_i(v_j)) = \det(A)\delta_{ij}$$

Proof. If $j \neq i$, then $A_i(v_j)$ has two equal columns: v_j appears in both the j-th and the i-th column and $\det(A_i(v_j)) = 0$

If $j = i$, then $A_i(v_i) = A$, so $\det(A_i(v_i)) = \det(A)$. $\square$

Corollary 8.22. $A(\mathrm{Cof}(A))^t = (\mathrm{Cof}(A))^t A = \det(A)I$. *If* $\det(A) \neq 0$ *then* $A^{-1} = \frac{1}{\det(A)}(\mathrm{Cof}(A))^t$.

Proof. By Theorem 8.19, $((\mathrm{Cof}(A))^t A)_{ij} = \sum_{k=1}^{n} (-1)^{i+k} A_{ki} a_{kj} = \det(A_i(v_j)) = \det(A)\delta_{ij}$. $\square$

Example 8.23. If $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ then $\det A = ad - bc$. If $ad - bc \neq 0$, then A is invertible and $A^{-1} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$